

O papel da qualidade da informação nos sistemas de informação¹.

João Nuno Oliveira

Luís Alfredo Amaral

Departamento de Sistemas de Informação

Universidade do Minho

Azurém

4800 GUIMARÃES

+ 351 53 510259

+ 351 53 510250

jnuno@dsi.uminho.pt

amaral@dsi.uminho.pt

<http://www.dsi.uminho.pt>

Resumo

Os problemas com a qualidade da informação são sentidos de forma rotineira por todas as organizações, com diferentes níveis de gravidade e de prejuízo. O impacto negativo traduz-se em custos desnecessários, em processos de decisão afectados ou na perda de confiança dos clientes. Evidências apontam que o problema da qualidade dos dados deve ser tratado com formalismo e deve seguir o caminho que a qualidade já tomou na indústria.

Ao encarar a informação como um produto, estamos a considerar a existência de uma economia da informação, na qual consumidores, fornecedores e produtores de informação desempenham um papel no processo de criação de valor. Um sistema de informação, à luz deste enquadramento, não é mais do que um processo de transformação de informação, algo semelhante a uma fábrica, cujo sucesso depende do grau de satisfação dos consumidores dessa informação.

Este artigo procurará abordar os sistemas de informação pelo conceito da qualidade, nomeadamente pelos princípios e métodos associados a um conceito já amplamente debatido noutras áreas. Discutirá o que é a qualidade da informação e as suas implicações na gestão da informação. Terminará com a definição de um modelo para sistemas de garantia da qualidade da informação, nomeadamente ao definir o papel do gestor do recurso informação e ao identificar um conjunto de actividades chave da função qualidade.

¹ Comunicação proferida na "Conferência Especializada em Sistemas E Tecnologias de Informação", que decorreu nos dias 4 e 5 de Outubro de 1999, nas instalações da Universidade Católica Portuguesa, em Lisboa

1. A INFORMAÇÃO COMO UM RECURSO

Nenhuma outra tecnologia revolucionou o mundo como a tecnologia da informação. A era da informação parece ter condenado legiões de cidadãos a depender de algo que é intangível, inodoro, invisível e de difícil quantificação ou valorização. A esta visão pessimista e negativista opõe-se uma outra, produto das mesmas sementes, de uma sociedade mais produtiva e mais cuidadosa para com os seus, ao disponibilizar o que mais de essencial existe para a nossa actuação: a informação. Algures no meio está a verdade de cada um.

Ao emergir destas especulações, vemos um conjunto de ferramentas que possibilitam o acesso cada vez mais rápido e eficiente a grandes volumes de informação em vários formatos, como a imagem ou o vídeo, e não limitado por questões geográficas ou temporais. São colocadas ao dispor das organizações e dos indivíduos ferramentas para que os seus objectivos sejam alcançados da forma mais eficiente e eficaz possível.

A informação é hoje encarada como um recurso essencial para as organizações (Amaral, L., 1994). Não que a informação seja uma novidade, até porque faz parte da própria essência das organizações, mas só nos últimos anos é que se reconheceu a sua importância estratégica. O seu reconhecimento como um recurso significa que esta terá que ser gerida como tal, que será necessária a existência de uma estrutura capaz de assegurar que a informação esteja disponível no momento, na forma e na quantidade desejável para os seus consumidores, ou seja, que tenha qualidade.

2. QUALIDADE DA INFORMAÇÃO

2.1. Dimensões da qualidade da informação

Todas as decisões que se tomam são baseadas num conjunto de informação que está disponível no processo de tomada de decisão, algo que faz depender a decisão final das características da informação que lhe serviu de *input*. Informação com insuficiente qualidade não produzirá uma decisão adequada que, quando aplicada, produza os resultados esperados.

Juran et al. (1993) definem qualidade como “*fitness for use*”. Apesar de muitas outras definições existirem, esta possui a vantagem de ser simples e explícita num aspecto extremamente importante da qualidade, que é o facto da qualidade não existir por si só nas características intrínsecas dos “objectos”, mas na utilização ou aplicação desses objectos. Assim como a qualidade de um carro só pode ser sentida por quem o usa, a qualidade da informação só pode ser avaliada por quem a consome, e só terá qualidade se quem a consome a considerar como apropriada para as necessidades em causa. E se facilmente encontramos características num carro, como a segurança, o consumo ou o desempenho, também na informação poderemos encontrar características determinantes para o seu “*fitness for use*”

Naturalmente, quando se fala em qualidade da informação, o termo “correção” é o primeiro a surgir, associando qualidade à ausência de erros na informação, embora até seja discutível o significado de “erro”. Para McKinnon e Bruns (1992), a informação para os executivos deverá ser correcta, relevante e actual. Wang et al. (1994) mostram-nos como as características da qualidade da informação vão para além da correção, ao considerarem aspectos como a apresentação, o acesso e a relevância como componentes importantes (quadro 1). Um conjunto de dados pode estar correcto, mas a sua apresentação dificultar a sua compreensão ou não foi disponibilizado em tempo útil. Estes aspectos diminuem a qualidade da informação junto do consumidor.

Dimensões	Características
Intrínseca	Correcção, objectividade, reputação e veracidade
Contextual	Valor acrescentado, relevância, disponibilidade, completa e volume apropriado
Representacional	Interpretabilidade, compreensão, consistência e concisa
Acessibilidade	Acessibilidade e segurança

Quadro 1 – Dimensões da qualidade da informação. Adaptado de (Wang, R. et al, 1994)

Wang et al. (1994) agruparam um conjunto de quinze características em quatro grupos:

Intrínseca: Características intrínsecas dos dados, independentes da sua aplicação

Contextual: Características dependentes do contexto de utilização dos dados

Representacional: Características derivadas da forma como a informação é apresentada

Acessibilidade: Aspectos relativos ao acesso e à segurança dos dados

Não existe unanimidade sobre as características e o significado das mesmas. Em primeiro lugar, as características existem aos olhos dos consumidores da informação, logo o conjunto de características variará de acordo com o contexto. Por outro lado, o significado das características também é objecto de discussão. Podemos encontrar alternativas às dimensões e características descritas em (Wand, Y. et al., 1996) e (Redman, T., 1996). Surge como a atitude mais correcta a definição, num dado contexto de utilização de informação, do conjunto de características e respectivos significados relevantes para a comunidade interveniente nesse contexto, e sobre tal conjunto abordar o problema relacionado com a qualidade da informação.²

Interessante também a associação de métricas a essas características. Por exemplo, Redman (1996) define como métrica para a *correção* o número de registos com pelo menos um erro sobre o total de registos, ou o total de campos errados sobre o número total de campos (número de registos vezes o número de campos de cada registo). Barquin et al. (1997) apresentam um conjunto de métricas para as “suas” características da qualidade da informação, como, por exemplo, a percentagem de regras de integridade suportadas para a característica “relacionabilidade”, definida como “a coerência lógica que permite a correlação relacional com outros dados similares”.

Até este momento nenhuma distinção tem sido feita entre “dados” e “informação”. É normal chamar *dados* a conhecimento no seu estado primário e independente da utilização desse conhecimento, como os dados que caracterizam os assinantes de um operador de comunicações móveis. *Informação* serão esses dados tratados, por um qualquer processo, para serem consumidos por alguém, como a lista dos assinantes com facturação superior a um dado valor. Um terceiro nível, o do *conhecimento*, representa a capacidade para utilizar a informação (Watson, R., 1999). Conhecimento significa que alguém pode interpretar a informação, utilizá-la nos seus processos de decisão e identificar a informação que será útil para essa decisão.

Para muitos autores, o tema da qualidade da informação é insensível à dicotomia dados/informação. O que são dados para uns será informação para outros, pelo que a avaliação da qualidade se fará do mesmo modo. Segundo outros autores, enquanto que os dados são independentes

² Um esforço de normalização interessante para o tema da qualidade dos dados pode ser encontrado no pré-standard europeu prENV 12656, “Geographic Information – Data description – Quality”, do CEN TC 287, para os sistemas de informação geográficos.

do contexto da sua utilização, quase ausentes de significado, a informação presuppõe um consumidor num dado contexto. No entanto, tal distinção não parece ter qualquer impacto na gestão da qualidade da informação.

2.2. A necessidade de gerir a qualidade da informação

Toda a literatura sobre qualidade foca o aspecto mais sensível para a nossa sociedade, o lucro. Toda e qualquer organização tem no resultado do seu exercício o indicador do seu sucesso, e, de entre alguns factores que contribuem para esse exercício, a qualidade ocupa um lugar de destaque. A aplicação prática do conceito da qualidade traduzir-se-á em aumentos nas vendas e redução de custos. A não-qualidade traduz-se invariavelmente no contrário, na redução das vendas e no aumento dos custos. A não-qualidade da informação não é diferente.

Várias evidências apontam para os custos directos ou indirectos derivados da insuficiente qualidade dos dados. Sente-se quotidianamente, quando se é destinatário de informação bancária incorrecta, quando vários trabalhadores têm como função resolver problemas de entregas e facturação deficientes, ou quando uma reserva de hotel afinal não foi feita.

Segundo Albrecth (1999), e muitos outros relatos podem ser encontrados em (Redman, T., 1996), (Strong, D., et al, 1997a), (Baer, T., 1995) ou (Strong, D., et al., 1997b):

- 120.000 cidadãos americanos morrem por ano devido a erros nos diagnósticos, segundo a American Medical Association
- Scanners em estabelecimentos comerciais apresentam um erro na leitura dos preços entre 1% e 3%. Apesar de aparentemente reduzida, a sua multiplicação pelas grandes quantidades dos produtos envolvidos, eleva significativamente o custo desse erro.

As consequências da falta de qualidade na informação são várias: custos acrescidos devido ao impacto que causam e os custos da sua reparação, a perda da confiança dos clientes, processos de tomada de decisão afectados, motivação das equipas diminuída ou processos de reestruturação organizacional, como a certificação ou a reengenharia, limitados pelo acesso a informação “utilizável”. A qualidade dos dados é citada frequentemente como um dos principais obstáculos das *data warehouses* (Barquin, R., et al., 1997).

Porquê então o relativo anonimato desta situação? Provavelmente porque só agora se atingiu um estado na nossa era da informação que tenha despoletado este problema. Actualmente, as necessidades operacionais de informação parecem estar satisfatoriamente suportadas. As transacções comerciais, monetárias ou logísticas possuem um elevado suporte informático, o que conduz a uma elevada taxa de recolha de informação com um custo marginal reduzido. As preocupações anteriormente centradas na quantidade de informação e no suporte aos processos operacionais conduziram a um cenário de informação em excesso, redundante e de difícil conciliação. As empresas são globais, a administração pública encontra-se em processo de informatização avançado, tudo parece ser digital. Aprende-se e descobre-se agora o que fazer com tanta informação disponível, potenciando sistemas de apoio à decisão, *data warehousing* ou *data mining*. Naturalmente, e por necessidade, a aposta actual vai no sentido da qualidade, para que seja possível produzir informação que seja útil e relevante para a gestão.

Por outro lado, lidar com a qualidade da informação implica uma mudança de atitude para com este problema. Não é mais um problema de tecnologia mas sim um problema de gestão. A tendência de considerar a informação como algo que está nas bases de dados, e por conseguinte da responsabilidade de um hipotético departamento de informática, terá que dar lugar ao reconhecimento da informação como um recurso e a este atribuir uma gestão adequada, com uma missão, um plano e recursos para a sua execução. Esta temática mais interessante fica quando se é confrontado com a existência de

organizações cujo produto que vendem é informação. As suas “linhas de produção” não escoam calçado ou placas de rede, mas informação. Qualidade do produto aqui incluirá necessariamente a qualidade da informação produzida. Exemplos desses cuidados são as empresas que produzem estudos de mercado, estudos de índole social ou económica, auditorias ambientais ou informação geográfica.

Apesar de a informação não ser um recurso como os tradicionais, dada a sua intangibilidade ou a ausência de unidades e de contornos físicos, tal não deve constituir uma desculpa para não tratar formalmente a gestão desse recurso.

3. A INFORMAÇÃO NAS ORGANIZAÇÕES

Qualidade tem sido um conceito amplamente debatido e aplicado na prática em vários campos, como na saúde, na indústria ou no software. No que concerne a informação, reconhecido como está a sua importância, a aplicação do conceito de qualidade é possível e produz mais valias no contexto da utilização de informação. Comparar um sistema de informação a uma indústria é estabelecer um paralelo útil para a aplicação das técnicas e das ferramentas da qualidade no campo da gestão da informação. Por outro lado, lidar com os problemas da qualidade da informação não constitui uma novidade, pelo que algumas técnicas ou abordagens já foram identificadas e discutidas. Esta secção vai apresentar em primeiro lugar o conceito de “fábrica de informação”, seguindo-se a apresentação de quatro abordagens para o controlo e a garantia da qualidade da informação.

3.1. As fábricas de informação

Um SI pode ser considerado como o conjunto gerido de recursos humanos e materiais, destinados a realizar as actividades de adquirir, armazenar, processar e difundir informação, quer estejam ou não envolvidos computadores. A sua missão é a de fornecer informação com qualidade a quem dela necessita (Spewak, S., 1993), a agentes internos ou externos à organização em causa.

DeLone e McLean (1992) consideraram a qualidade da informação como uma dimensão do sucesso de SI. Apesar do modelo para avaliar o sucesso de um SI por eles proposto ser objecto de críticas (Ballantine, J., et al, 1996), é no entanto reconhecido como um trabalho de referência. Deste modo, a qualidade da informação é um elemento a considerar na avaliação da qualidade de SI, a par de outros elementos, como seja a qualidade do *software* ou do serviço de suporte.

A comparação entre um SI e uma fábrica pode ser encontrada em (Wang, R., 1998). Assim como uma fábrica transforma as matérias-primas que adquire em produtos acabados e prontos a serem enviados para os clientes, um SI adquire informação (ou dados, segundo alguns) e os transforma em mais informação, informação esta fornecida aos consumidores. Este paralelismo tem a grande vantagem de estabelecer a ponte entre os conceitos da qualidade na indústria e os conceitos da qualidade na indústria da informação, apesar do perigo que é tratar a informação como um produto com instanciação física.

Para poder actuar sobre qualidade da informação, é necessário compreender o ciclo de vida da informação. Basicamente, a informação é recolhida, armazenada e mais tarde extraída para ser consumida. Dois ciclos podem ser identificados, o ciclo de aquisição e o de utilização (Redman, T., 1996) (Levitin, A., et al, 1993) (Fig. 1). O ciclo de aquisição inicia-se com a definição do modelo conceptual de dados ou a visão do mundo real que deve ser recolhida. Prossegue com a implementação do modelo, a aquisição de valores e o seu armazenamento. O ciclo de utilização começa com a definição do subconjunto de dados do modelo conceptual, a sua extracção, manipulação e apresentação. A última actividade é a sua utilização pelo consumidor. Ambos os ciclos são executados indefinidamente.

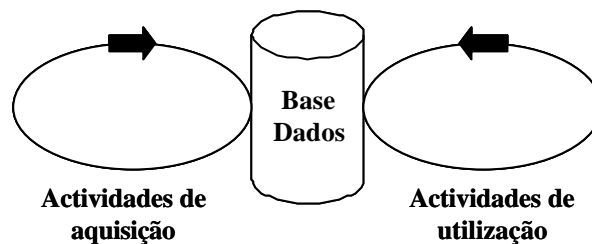


Fig.1 – Ciclo de vida da informação. Adaptado de (Redman et al., 1996)

Em diversos pontos dos ciclos existem pontos de inspecção. Nesses momentos a informação deverá ser avaliada, ou seja, indicadores sobre a qualidade devem ser recolhidos e posteriormente analisados. Consoante os resultados, o processo voltará a um ponto a montante, ponto esse que será determinado pela gravidade das não-conformidades detectadas. Está implícito para este modelo que os requisitos da qualidade foram identificados e quantificados e que os métodos para recolha dos indicadores da qualidade estão definidos.

Quatro intervenientes neste processo podem ser identificados (Wang, R., 1998): os *fornecedores*, responsáveis por criar ou recolher a informação, os *consumidores*, os *produtores*, aqueles que são responsáveis pelo desenvolvimento, exploração e manutenção dos sistemas e dos dados armazenados, e os *gestores*, responsáveis últimos pela informação em todo o processo.

O facto de a informação não ser um recurso tradicional, com dimensões físicas, levanta um conjunto de questões. Pelo facto de não ter dimensões, é difícil avaliar o seu volume ou conseguir identificar a quantidade de informação adequada. A informação não se consome com o uso, é um recurso que não se esgota, pois pode ser utilizado indefinidamente, o que levanta problemas no controlo da sua utilização. A própria analogia entre os sistemas de informação e uma fábrica falha no facto da utilização da informação não ser determinística. Num processo fabril, a sequência das etapas de fabrico estão estabelecidas e a utilização do produto intermédio e do final estão definidas. Num sistema de informação, os utilizadores podem a qualquer momento recorrer à informação existente e aplicá-la de uma forma nunca antes imaginada.

3.2. Abordagens ao controlo e à garantia da qualidade dos dados

Pretende-se aqui apresentar quatro abordagens complementares ao controlo e garantia da qualidade da informação. Estas reflectem a prática corrente e são a detecção e correcção de erros, a análise do processo, metadados da qualidade dos dados e a qualidade total.

3.2.1. Detecção e correcção de erros

Esta técnica tem como alvo a característica “correcção” da informação. Consiste em analisar os dados presentes numa base de dados e detectar incorrecções, como valores nulos, informação duplicada ou errada. No entanto, a informação só estará correcta se reflectir a realidade, mas comparar com a realidade milhões de registos é impraticável. Por isso é necessário recorrer à recolha de amostras para posteriormente comparar essas amostras com os valores reais. Esta comparação fornecerá indicadores sobre a qualidade dos dados presentes na base de dados. A operação de limpeza consiste

na correcção da informação errada encontrada. Sendo uma característica intrínseca, a presença dos consumidores da informação não é necessária, no entanto, esta faz-se sentir porque são muitas vezes os consumidores os responsáveis pela detecção de informação incorrecta.

A origem das incorrecções dos dados pode estar localizada na recolha da informação ou na introdução dos dados no sistema. Alguns erros podem ser evitados com o recurso às restrições de domínio e de integridade presentes na maioria dos sistemas de gestão de bases de dados, mas mesmo a identificação destas restrições não será nunca completa para um dado contexto. Por outro lado, sistemas de recolha automática de informação reduzem em muito os erros de digitação.

A existência de várias ferramentas informáticas simplifica esta tarefa. Estas podem pesquisar volumosas bases de dados para a detecção de possíveis erros, como duplicação de informação, informação inexistente, validação de restrições, valores “anormais, ou tratamento específico de nomes e moradas de clientes.

3.2.2. Analise do processo

Corrigir os erros contidos numa base de dados não é suficiente para assegurar a médio e longo prazo uma base de dados isenta de deficiências. Uma limpeza de uma base de dados não impede que surjam mais tarde mais erros e que esta não volte ao seu estado anterior. A razão disto prende-se com o facto de que o processo responsável pela recolha e introdução dos dados não sofreu qualquer alteração. É necessário analisar o processo e detectar os pontos nos quais os erros são introduzidos. É a resposta à actuação de inspecção da abordagem anterior, com a ênfase a ser agora colocado na prevenção dos erros.

Em Redman (1996) e em Huh et al. (1990) podemos encontrar um sistema de “*data tracking*”, que consiste em marcar certos registos de dados no momento da sua recolha e observá-los ao longo do seu percurso. Esta técnica permite observar as alterações que os dados sofrem ao longo do ciclo de vida, para desta forma detectar os pontos nos quais são introduzidos erros ou distorções. Ainda em Redman (1996) encontramos uma representação pictórica para o processo de recolha e introdução de dados.

Técnicas comuns na área da qualidade, como as cartas de controlo ou os diagramas de Pareto, são utilizadas para visualizar a evolução dos defeitos encontrados na análise de tabelas, registos ou mesmo campos. Não só visualizar, mas também monitorar a evolução das ocorrências. Para um processo estabelecido de transformação de informação, e por estabelecido entenda-se formalizado e a operar em pleno, é possível colocá-lo sob controlo estatístico, para desta forma observar a evolução dos indicadores da qualidade.

3.2.3. Metadados da qualidade dos dados

Quando um consumidor de informação utiliza um conjunto de tabelas de uma base de dados para pesquisar informação de que necessita, não tem ao seu dispor informação alguma sobre a qualidade dos dados que tem à sua frente. Uma forma de colmatar essa lacuna consiste na inclusão de informação sobre a qualidade desses dados, recorrendo a campos ou a tabelas adicionais.

Na figura 2 pode ver-se um modelo de informação com a inclusão de elementos ou de indicadores da qualidade dessa informação.

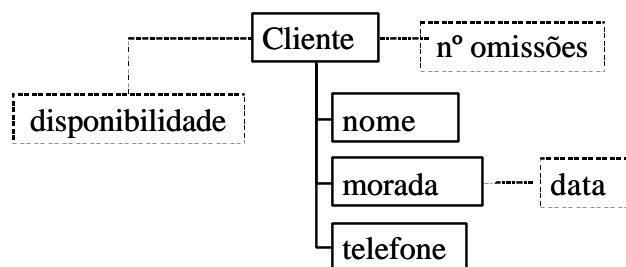


Fig. 2 – Modelo de dados com indicadores da qualidade

A “data” indica a idade da morada, algo que pode ser utilizado na avaliação da sua actualidade. O atributo “nº de omissões” será, uma vez implementado, utilizado para indicar o número de registos omissos, ou seja, nos quais falta preencher um ou mais atributos. “Disponibilidade” pode dar a indicação do intervalo de tempo entre a recolha da informação sobre um cliente e o momento em que a informação é introduzida no sistema. Um tempo de disponibilização elevado indica que a base de dados não espelha os cliente actuais.

3.2.4. Qualidade total

O “total quality management” (TQM) representa a abordagem mais organizacional ao problema da gestão da qualidade. É antes de mais uma filosofia ou atitude baseada no princípio da melhoria contínua, na crença de que é sempre possível fazer mais e melhor. É uma abordagem disciplinada que integra técnicas de gestão com ferramentas da qualidade.

Os conceitos associados, bem como as técnicas e as ferramentas, foram já muito explorados em campos como na indústria. A sua importação para o campo da qualidade da informação é tentadora, desde que salvaguardadas as diferenças entre o produto “informação” e os tradicionais produtos de manufactura.

Central na actuação da qualidade total está o ciclo Deming para a melhoria contínua: “planear-fazer-avaliar-actuar”, que consiste em estabelecer objectivos, recolher medidas de desempenho actual, comparar os indicadores de desempenho com os objectivos traçados e actuar, alterando o sistema para diminuir a diferença de resultados obtida. Em Wang (1998) podemos encontrar o “total data quality management”, ou seja, a aplicação do conceito de TQM à qualidade dos dados. A abordagem proposta assenta em três pilares básicos: o produto, o ciclo de melhoria contínua e a gestão do processo.

O produto informação deve ser descrito, realizando para tal a identificação das suas características e dos seus requisitos de qualidade, e a definição do sistema que irá produzir o produto. O ciclo de melhoria é uma adaptação do ciclo de Deming e é composto por quatro actividades (Fig. 3). Ao nível da estação do processo, o autor preconiza a constituição de uma equipa responsável pelo produto e composta por representantes dos quatro intervenientes da já apresentada fábrica de informação.

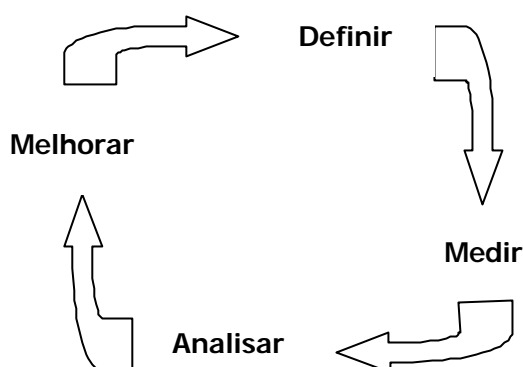


Fig. 3 – Ciclo de melhoria contínua para a qualidade da informação.

Adaptado de (Wang, R., 1996)

Abordagem semelhante pode ser encontrada em (Willshire, M., 1997).

Para concluir esta secção, é importante afirmar que as técnicas utilizadas ou até mesmo os recursos empregues na resolução dos problemas derivados da qualidade da informação dependem da gravidade que estes assumem nas organizações. Para algumas organizações, um bom conjunto de restrições na base de dados serão suficientes para alcançar os níveis desejados de qualidade. Pelo contrário, outras instituições necessitam de sistemas mais elaborados, como por exemplo no campo da saúde, no qual estão em risco seres humanos. Mesmo dentro das organizações, a informação não apresenta toda ela as mesmas características. Alguma informação poderá necessitar de um controlo mais apertado do que outra. É imperioso adequar as técnicas e os esforços dispendidos aos níveis de qualidade desejados na organização.

4. UM MODELO PARA A GARANTIA DA QUALIDADE DA INFORMAÇÃO

Ao designar a informação como um recurso e ao atribuir-lhe os recursos necessários para que a informação consumida tenha as características adequadas à sua utilização, ou seja, tenha qualidade, estamos a criar um sistema de garantia da qualidade da informação. Tal sistema deverá assegurar que a função qualidade está alinhada com a estratégia da empresa, está instalada e que obtém os resultados pretendidos. Segundo Ballou e Tayi (1996), a gestão da qualidade dos dados “envolve a especificação de políticas, a identificação de técnicas e a utilização de procedimentos para assegurar que os dados da organização possuem o nível de qualidade adequado para a sua utilização actual e futura”

Como o modelo para o “*total quality management*” de Oakland (1993), também o “*total quality data management*” deverá estar centrado nas *cadeias de informação* (Redman, T., 1996), ou seja, nas sequências de transformações que conduzem a informação dos fornecedores até aos consumidores. Esta decisão deve ser acompanhada pelo *comprometimento* da gestão ou mesmo da organização para com a qualidade da informação e a melhoria contínua, pela criação de uma *cultura* da qualidade e pelo fomento da *comunicação* entre departamentos ou áreas da organização, dada a transversabilidade da informação. Os pilares básicos para levar o barco da qualidade dos dados a bom porto são o *trabalho de equipa*, o *sistema* da garantia da qualidade e as *técnicas* da qualidade.

4.1. A gestão da função qualidade da informação

Será um objectivo do gestor da qualidade da informação assegurar que todas as actividades ou processos que de alguma forma podem influenciar a qualidade da informação estão identificados e controlados.

A nomeação de um gestor para recurso informação é necessária por questões organizacionais, ao ser-lhe entregue a responsabilidade implementar a política da qualidade da informação. O reconhecimento deste cargo representa a evolução que a importância deste recurso tem tido. Se se começar pelos sistemas de ficheiros, vê-se que a noção de informação dilui-se no conceito de aplicação, no qual os dados não existem por si só, mas apenas associados às aplicações informáticas que os envolvem. Com o aparecimento das bases de dados, a informação separa-se das aplicações, e surge o *gestor das bases de dados*. No entanto, a este apenas são atribuídas funções relacionados com os acessos e a segurança da base de dados, ao mesmo tempo que questões como a partilha de informação entre departamentos e a definição de um modelo conceptual de informação coerente e global à organização são ignoradas. O *gestor dos dados* procura colmatar as lacunas evidenciadas, apesar de a distinção entre o que é administração de dados e administração de bases de dados não ser clara (Guynes, C., et al, 1996). O primeiro assume funções mais de gestão, enquanto que ao segundo são atribuídas funções iminentemente técnicas. O *administrador do recurso informação* será o passo seguinte, como sinal do reconhecimento da informação como um recurso e do impacto desta na organização. A *gestão da informação* será o garante do correcto enquadramento entre consumidores, produtores e fornecedores de informação, para garantir as necessidades actuais e futuras da organização.

As responsabilidades atribuídas a um gestor da informação são, entre outras:

- Reconhecer a natureza da informação e a importância desta no negócio
- Definir, planear e aplicar as políticas que asseguram a qualidade da informação
- Encontrar o ponto de equilíbrio entre requisitos ou necessidades contraditórios
- Gerir em termos económicos a função qualidade da informação

Deve existir na base de um sistema de garantia da qualidade da informação a política da empresa sobre esta matéria, algo que sirva como referencial para todas as actividades e decisões, e que traduza a visão e o comprometimento do mais alto nível da gestão. Tal declaração deve ser divulgada por todos na organização e funcionará como elemento motivador. Wang et al. (1995) reconhecem que este aspecto se encontra pouco explorado. O que é uma política de qualidade da informação, que conteúdo e que forma deve ter, são ainda questões em aberto e que só a sua prática poderá conduzir a um maior esclarecimento. Em Redman (1996) pode ver-se o assunto tratado com algum detalhe.

4.2. Actividades da função qualidade

Embora a qualidade da informação esteja distribuída por diferentes áreas e sectores de uma organização, a função qualidade determina um conjunto de actividades necessárias para cumprir a sua missão. É aqui descrito um conjunto de actividades que necessariamente farão parte da função qualidade da informação. São elas a manutenção de um inventário do recurso informação, a criação do suporte normativo, a identificação e correcção de deficiências, a formação, o lançamento e o controlo de projectos de melhoria e o controlo do desenvolvimento de produtos de informação.

4.2.1. Manter um inventário do recurso informação

Será uma inutilidade iniciar projectos de melhoria da qualidade da informação se não se souber em que estado esse recurso se encontra. Que dados existem numa organização, quem e quando os produz, quem e com que finalidade os consome, são perguntas a que um inventário deverá responder. À semelhança de outros recursos, as organizações precisam de saber o que está ao seu dispor em termos do recurso informação. Trata-se da constituição de um inventário. Esta necessidade foi, por exemplo, reconhecida por Willshire (1997), que num processo de melhoria da qualidade da informação, explicita a necessidade de construir essa visão panorâmica, recorrendo para o efeito a técnicas de análise comuns, como a modelação de dados e aos diagramas de fluxo.

Do inventário da informação deve fazer parte o dicionário de dados, aquilo que dá significado aos dados. Uma das principais origens apontadas aos problemas que surgem no momento da utilização de informação é a separação semântica entre quem produz e quem consome informação. Como já referenciado, dada a natureza dinâmica da informação, não é possível antever os cenários nos quais a informação será utilizada, o que aliado à partilha descoordenada de informação, se traduz na perda do significado atribuído a um dado conjunto de dados. É comum encontrar significados diferentes para campos com nomes iguais. Mesmo conceitos que se julgam universais nas organizações assumem significados diferentes nos vários departamentos. Dados sobre os dados, metadados, devem ser mantidos e acessíveis a toda a organização (Witt, G., 1998).

Outra questão se pode levantar. Que informação ficará sob a responsabilidade do gestor da informação? Apenas a informação de índole operacional e tradicionalmente já armazenada nas bases de dados ou também textos ou desenhos? De qualquer forma, esta fronteira deverá ficar bem delineada para que a função qualidade tenha uma base sólida para operar.

4.2.2. Criar o suporte normativo

Quando se fala de normas, procedimentos ou de regras, rapidamente o fantasma da burocracia se levanta e começa a saltar de alegria. Por outro lado, nenhuma organização assegura consistência e coerência sem um conjunto de regras mínimo. Claramente, a existência de regras ou de normas é um produto do equilíbrio entre a burocracia e a anarquia. Uma organização deve limitar o seu esquema normativo ao que considera essencial e lhe garante o controlo do seu funcionamento, para que possa oferecer qualidade no que faz. Além disso, qualquer *standard* ou norma deve ser fruto de uma necessidade sentida no terreno.

Cabe à função qualidade definir o conjunto de políticas, procedimentos e de normas que considere necessárias para o bom desempenho do sistema de garantia da qualidade da informação e conducentes a uma informação de qualidade³.

Como exemplo, podem ser referenciadas as políticas e normas para questões como o acesso à informação, segurança dos dados, métodos para recolha de informação, instrumentos de recolha e análise de informação, processos de migração, definição de requisitos de informação, entre outros.

³ A instituição americana *Environmental Protection Agency* constitui um exemplo muito interessante de um sistema documentado para garantia da qualidade da informação (<http://www.epa.org> ou <http://es.epa.gov/ncercqa/qa/index.html>)

4.2.3. Identificação e correcção de deficiências

As deficiências na informação que afectam a sua qualidade podem ter origem em qualquer ponto do processo. Paralelamente, oportunidades para melhorar podem ser identificadas também em qualquer ponto do mesmo processo. É missão da função qualidade criar condições para as deficiências identificadas dêem origem a medidas correctivas e as oportunidades em medidas perfectivas. Com as primeiras procura-se evitar a reincidência do defeito e com as segundas procura-se uma melhoria na eficácia e na produtividade.

O tratamento de uma deficiência deve seguir um conjunto de passos, nomeadamente:

- Identificação – Uma vez identificada uma situação de não-conformidade, esta deve ser documentada.
- Análise de impacto – Avaliar o impacto na organização da deficiência detectada. O resultado deverá indicar a gravidade da deficiência.
- Determinar as causas – Análise e identificação das causas para a deficiência.
- Definir acções correctivas – Para que deficiências do mesmo tipo não se repitam, medidas devem ser tomadas. Tais medidas podem dar origem a alterações nos procedimentos, no processo de transformação ou nos processos de recolha de dados.
- Controlo dos resultados – A aplicação das medidas deve ser objecto de controlo e de verificação. O resultado das medidas depois de aplicadas deve ser observado para verificar o impacto desejado.

4.2.4. Formação

À função qualidade são atribuídas responsabilidades na área da formação e da sensibilização. Uma vez que a qualidade é uma preocupação de todos e de cada um, há que adequar as pessoas à função ou tarefa que executam, dotando-as dos conhecimentos e das técnicas mais apropriados.

A formação na área específica da qualidade dos dados é escassa. Uma pesquisa em vários livros sobre bases de dados revela a quase inexistência de informação sobre esse tema, à excepção de incursões na questão das dimensões da qualidade e também nos terrenos habituais das bases de dados, como a integridade dos dados, o controlo de acessos, a gestão da concorrência e a recuperação de bases de dados (Mathieu, R., et al, 1998).

Por outro lado, e apesar do reconhecimento da importância do tema, a sensibilidade das organizações para o problema é diminuta, o que afecta necessariamente a actuação neste aspecto da formação.

Esta preocupação em manter os “skills” apropriados estende-se a todos os indivíduos que de alguma forma se encontram envolvidos em actividades que podem afectar a qualidade da informação, sejam eles parte do grupo dos fornecedores, consumidores, produtores ou dos gestores.

4.2.5. Projectos de melhoria

Qualidade é essencialmente um problema de atitude, e parte dessa atitude é a aceitação de que se pode fazer sempre mais pela qualidade em termos economicamente aceitáveis. Aceitar que se está no topo e que nada mais pode ser feito para melhorar é o primeiro passo para perder a liderança ou as

esperanças de a alcançar. É este tipo de mensagem que deve circular nas organizações e é este tipo de atitude que qualquer indivíduo deve ter.

Uma atitude participativa será outra trave mestra da qualidade. A qualidade não é da responsabilidade do departamento da qualidade mas de todos no seio da organização. Se se reconhece competência nos indivíduos envolvidos, então estes constituem uma fonte de inovação e de informação não negligenciável para a melhoria contínua, pelo que é necessário envolvê-los activamente na procura das melhores soluções.

É da responsabilidade da função qualidade criar condições para que a participação seja uma realidade e que as propostas tenham um encaminhamento apropriado, quer sejam aceites ou rejeitadas. Os célebres círculos da qualidade japoneses são um exemplo disso. Procura-se que se criem equipas voluntárias de trabalhadores com o propósito de resolver um problema, apontando soluções ou identificando inovações no processo ou no produto. Procura-se mesmo que sejam os próprios a aplicar as acções correctivas ou perfectivas. À função qualidade cabe a tarefa de criar condições para que esses grupos se formem, obtenham os recursos necessários e que o seu trabalho seja observado e controlado para garantir a compatibilidade organizacional.

Quer seja de uma forma espontânea ou centralizada, cabe à função qualidade a identificação dos projectos de melhoria, a realização de análises custo-benefício das diversas propostas, a calendarização das acções, o controlo da execução das acções e a análise de resultados.

O primeiro projecto é determinante do sucesso de uma política a longo prazo para a qualidade da informação. Deve ser um projecto equilibrado, ao procurar um problema com uma complexidade tal que garanta alcançar sucesso, mas que ao mesmo tempo possua uma dimensão suficiente para que o sucesso se faça sentir e seja visível. Deve ser um projecto sério, isto é, que corresponda a um real problema e que a equipa tenha os meios para actuar no terreno. Se este primeiro projecto é um sucesso outros seguir-se-ão, e desta forma a qualidade da informação vai adquirindo o seu espaço na organização.

4.2.6. Concepção e controlo do desenvolvimento de um produto de informação .

Os consumidores não querem sistemas de informação, mas sim informação com qualidade. Um SI é concebido e construído com o propósito de fornecer informação com qualidade aos seus utilizadores.

É reconhecido, provavelmente desde o seu início, que a actividade de desenvolvimento de SIs não produz os resultados esperados. Os constantes atrasos no desenvolvimento e o reduzido nível de satisfação dos utilizadores são a norma e não a excepção. Parte dos insucessos poderão ser atribuídos à não satisfação dos requisitos de qualidade da informação dos utilizadores. E, de facto, são raros os produtores de informação que procuram iniciar o desenvolvimento de um sistema pelo levantamento das características desejadas para a informação, uma actividade quase de “marketing”. Pelo contrário, assiste-se à imediata análise dos *inputs* de informação e das regras de processamento que conduzem os dados até à base de dados. Aparentemente, o ciclo de utilização da informação, a extracção de informação, é um resultado natural e simples, fruto da aquisição e do armazenamento dos dados.

A abordagem descrita levanta vários problemas:

- A maioria dos *inputs*, sejam eles formulários ou informação electrónica, não são adequados. Apresentam incorrecções, lacunas e são frequentemente utilizados de outras formas que não as inicialmente planeadas. É preciso ver que estes elementos foram estudados para servir necessidades operacionais.

- O armazenamento de dados, nomeadamente o modelo conceptual, está muito mais vocacionado para as necessidades de aquisição do que as de extracção.
- A utilização de informação surge aparentemente como assegurada, pois apenas corresponde a diferentes combinações dos dados recolhidos.
- É recolhida e armazenada informação que não tem qualquer utilidade.
- A questão da qualidade resume-se à definição de um conjunto de restrições aplicadas como filtros na entrada dos dados e à responsabilização do utilizador, aqui no papel de fornecedor, por toda a informação que escapou às restrições colocadas. A responsabilidade pela má qualidade dos dados parece repartida entre produtores e fornecedores.

Para evitar estas questões é preciso mudar de atitude. Inevitavelmente, o desenvolvimento de um produto de informação, algo que seja fruto de uma necessidade de um ou de um conjunto de consumidores, começa pela definição das características desse produto e dos seus requisitos de qualidade.

Depois de definido o produto, passa-se à avaliação do estado do recurso informação para ver se o produto pretendido pode ser disponibilizado. Numa primeira instância, pode ser apenas necessário definir o ciclo de utilização de informação, que consiste, como já foi dito anteriormente, na definição do conjunto de dados necessário, na construção dessa “view”, e na extracção e posterior apresentação da informação. Se tal não for possível, será então necessário construir todo o caminho que se inicia nas fontes de informação, passa pela construção do ciclo de aquisição de dados, armazenamento e o posterior ciclo de utilização.

Abordagens ao desenvolvimento de SIs, como o SSADM (Downs, E., et al, 1992), distinguem claramente processos destinados à aquisição de informação e processos destinados à utilização da informação, mas falham ao não considerar explicitamente as necessidades de informação dos consumidores e a tradução das mesmas em requisitos da qualidade.

Para que a qualidade da informação seja salvaguardada é necessário que as actividades de desenvolvimento de um SI sejam capazes de traduzir os requisitos de informação dos consumidores em sistemas capazes de os satisfazer. O tratamento específico da qualidade da informação deve ser parte daquilo que é desenvolver um SI. Por outro lado, o desenvolvimento é uma actividade que deverá estar sob a alçada do gestor da informação.

4.2.7. Fomentar a utilização da informação

O gestor da informação deve ser detentor da visão do estado actual da informação mas também de uma visão estratégica, de uma visão que aponte os caminhos a seguir para que a informação preste o papel mais útil à organização, permitindo a própria evolução desta última. Partirá da função qualidade a iniciativa de fomentar junto da organização novas utilizações da informação. Ao gestor corresponderá a um papel de animador, ao procurar mostrar o que é possível fazer em termos de informação e de conhecimento e, em parceria com os demais intervenientes, evoluir a gestão do recurso. Esta atitude fará girar a “roda” da qualidade, movendo a informação e a sua gestão para níveis cada vez mais elevados. Quanto mais informação for utilizada, maior será a qualidade desta e maior será o efeito positivo na organização.

Orr (1998) afirma, baseado nos sistemas de controlo com realimentação, que a melhor forma de melhorar a qualidade da informação é aumentar e melhorar o seu uso. O consumidor é o único elemento capaz de identificar desvios entre a informação que recebe do sistema e a sua contrapartida no mundo real. Quanto mais severa for a utilização da informação, maior número de não-conformidades serão. Dessa visão dos sistemas de informação, Orr identificou um conjunto de regras para a qualidade da informação:

- Dados não utilizados não podem permanecer correctos muito tempo.
- A qualidade da informação é função do seu uso e não da sua recolha e armazenamento.
- A qualidade dos dados não será melhor do que a correspondente à sua utilização mais severa.
- Os problemas com a qualidade da informação tendem a agravar-se à medida que o sistema envelhece.
- Quanto menor for a frequência de alteração de um dado elemento, mais traumática é essa mudança.
- As leis da qualidade da informação aplicam-se aos dados e aos metadados.

Como conclusão desta secção, não será estranha alguma semelhança entre o que aqui foi dito e o esquema normativo das ISO9000. A justificação poder-se-á basear em dois pontos. Em primeiro lugar, é de qualidade que se tem vindo a falar e as normas ISO9000 são exactamente isso. Em segundo, as normas já provaram ser um auxiliar precioso na criação e na manutenção de sistemas de garantia da qualidade. O grande problema é que informação não é como os produtos “tradicionais”, para os quais a norma foi idealizada. Por exemplo, a informação não se compra nem se vende, o seu acesso é que pode ser adquirido, não se desgasta com o uso, o seu valor não é exacto, não se substitui informação e a sua replicação é ilimitada.

No entanto, as questões mantêm-se. Como pode uma organização que vende informação garantir que esse produto será entregue ao consumidor e que respeitará os requisitos impostos? Como pode uma organização garantir que utiliza informação com qualidade?

5. CONCLUSÕES

As empresas têm diferentes níveis de requisitos de qualidade na informação que utilizam, pelo que devem tomar medidas adequadas a esses níveis. Qualquer actuação para identificar, resolver e prevenir problemas na informação implica um custo que será compensado pela melhoria conseguida. Gerir a qualidade da informação será equilibrar o peso que se está disposto a suportar para obter o nível desejado de qualidade, mas procurar sempre fazer mais com menos.

Uma empresa deverá iniciar o caminho da garantia da qualidade da informação por questionar a existência de problemas relacionados com a natureza da informação que usa. Se houver o reconhecimento da importância que a qualidade dos dados desempenha no seu funcionamento, deverá então lançar um programa com as dimensões adequadas ao peso do seu problema. Esta adequação traduzir-se-á no âmbito das medidas a tomar, nos contornos do sistema a implementar e no valor dos recursos a adquirir.

Esta artigo procurou lançar as bases do que será um sistema de garantia da qualidade da informação, ao especificar um conjunto de princípios e de actividades que compõem esse sistema. É um trabalho que ainda se encontra incompleto, pois só a prática poderá colmatar as lacunas existentes. Esta preocupação com a gestão da qualidade enquadra-se num novo cenário as tarefas de desenvolvimento de sistemas de informação, ao lançar um conjunto de desafios com o objectivo de adequar a actividade aos objectivos da qualidade da informação.

Uma menção à actividade de planeamento de sistemas de informação é necessária. De uma forma clara, a política da qualidade da informação a adoptar é algo que tem a sua origem no topo da gestão, fruto dos processos de decisão estratégicos. No entanto, este tema encontra-se ausente da maior parte das abordagens para o planeamento de sistemas de informação, actividade que surge direccionada mais para a identificação da arquitectura do sistema de informação e para a identificação

das aplicações a instalar. Esta lacuna merece que este tema da qualidade da informação seja investigado no âmbito das actividades de planeamento de sistemas de informação.

Bibliografia

- Albrecht, K., *Information: The Next Quality Revolution*, in *Quality Digest*. 1999.
- Amaral, L., *PRAXIS - Um Referencial para o Planeamento de Sistemas de informação*, in *Departamento de Informática*. 1994, Universidade do Minho: Braga.
- Baer, T., *Mister Data Clean*, in *Computerworld*. 1995.
- Ballantine, J., et al., *The 3D Model of Information Systems Success: the Search for the Dependent Variable Continues*. Information Resources Management Journal, 1996. 9 (4): p. 5-14.
- Ballou, D. and G. Tayi. *Managerial Issues in Data Quality*. in *Information Quality*. 1996. Cambridge, Massachusetts, USA.
- Barquin, R. and H. Edelstein, *Planning and Designing the Data Warehouse*. 1997: Prentice Hall.
- DeLone, W. and E. McLean, *Information Systems Success: The quest for the dependent variable*, in *Information Systems Research*. 1992. p. 60-95.
- Downs, E., P. Clare, and I. Coe, *SSADM - Structure Systems Analysis and Design Method: Application and Context*. 2 ed. 1992: Prentice-Hall.
- Guynes, C. and M. Vanecek, *Critical Success Factors in Data Management*, in *Information and Management*. 1996. p. 201-209.
- Huh, Y., et al., *Data Quality*, in *Information and Software Technology*. 1990. p. 559-565.
- Juran, J. and F. Gryna, *Quality Planning and Analysis*. 3 ed. Industrial Engineering and Management Science, ed. McGraw-Hill. 1993.
- Levitin, A. and T. Redman, *A Model of the Data (Life) Cycles with Application to Quality*, in *Information and Software Technology*. 1993. p. 217-223.
- Mathieu, R. and O. Khalil, *Data Quality in the Database Systems Course*, in *Data Quality Journal*. 1998.
- McKinnon, S. and W. Bruns, *The Information Mosaic: How Managers Get The Information They Really Need*. 1992, Boston: Harvard Business School Press.
- Oakland, J., *Total Quality Management*. 2ª ed. 1993: Heinemann Professional Publishing Ltd.
- Orr, K., *Data Quality and Systems Theory*, in *Communications of the ACM*. 1998.
- Redman, T., *Data Quality for the Information Age*. 1996: Artech House Inc.
- Spewak, S., *Enterprise Architecture Planning - Developing a Blueprint for Data, Applications and Technology*. 1993: John Wiley & Sons.
- Strong, D., Y. Lee, and R. Wang, *10 Potholes in the Road to Information Quality*, in *Computer Magazine*. 1997.
- Strong, D., Y. Lee, and R. Wang, *Data Quality In Context*. Communications of the ACM, 1997. 40 (3): p. 103-110.
- Wand, Y. and R. Wang, *Anchoring Data Quality Dimensions in Ontological Foundations*, in *Communications of the ACM*. 1996. p. 86-95.
- Wang, R., *A Product Perspective on Total Data Quality Management*, in *Communications of the ACM*. 1998. p. 58-65.
- Wang, R., D. Strong, and L. Guarascio, *Data Consumers Perspectives of Data Quality*, . 1994, Total Data Quality Management Group - Massachusetts Institute of Technology.

- Wang, R., V. Storey, and C. Firth, *A Framework for Analysis of Data Quality Research*, in *IEEE Transactions on Knowledge and Data Engineering*. 1995. p. 623-640.
- Watson, R., *Data Management - Databases and Organization*. 2 ed. 1999: John Wiley & Sons.
- Willshire, M., *A Process for Improving Data Quality*, in *Data Quality*. 1997.
- Witt, G., *The Role of the Meta Data in Data Quality*, in *Journal of Data Warehousing*. 1998. p. 2-10.